UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/610,684	07/02/2003	Scott Shepard	02-4038	2914

25537 7590 01/24/2008

VERIZON
PATENT MANAGEMENT GROUP
1515 N. COURTHOUSE ROAD
SUITE 500
ARLINGTON, VA 22201-2909

EXAMINER	
SIEDLER, DOROTHY S	

ART UNIT	PAPER NUMBER
2626	

NOTIFICATION DATE	DELIVERY MODE
01/24/2008	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patents@verizon.com

Office Action Summary	Application No.	Applicant(s)	
	10/610,684	SHEPARD ET AL.	
	Examiner	Art Unit	
	Dorothy Sarah Siedler	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *15 November 2007*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-47* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-47* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Response to Arguments

1. Applicant's arguments filed November 15, 2007 have been fully considered but they are not persuasive. Applicant argues that, "Shiotani is a machine translation system, wherein a human being is not involved in the actual translating from one language to a different language. At most, a human being is involved in Shiotani in correcting the result of the previously-made translation, which was made by machine, but is not involved in the translating process itself" (Remarks page 16). Applicant continues by stating, "This is not describing a human translation process by which a human being translates from one human language to a different human language, as claimed. Rather, this is describing a "correcting" process by which a human being can make correction to a language IN THAT SAME LANGUAGE. Indeed, making any change (e.g. a correction) to words expressed in e.g., "language A" which remains in "language A" is plainly not a translating process." (Remarks page 16). However the examiner respectfully disagrees, and contends that the correction step taken by the user is in fact part of the translation process. The user observes the source utterance, then the target utterance (Figure 4(a) and 4(b)) and determines that the target utterance is incorrect, i.e. incorrectly translated, and provides the correct translation. This combination of human and machine for translating purposes is used to insure correct translation, especially between languages from different language families. Therefore

meets the limitation, "receiving translation made by the user", as recited in amended claim 1.

2. Applicant states that, 'it is noted that Schulz is limited to transcription and does not mention 'translating' or 'translation' even once" (Remarks page 17) as evidence that the combination of ***Shiotani*** in view of ***Schulz*** is improper; However the examiner respectfully disagrees. ***Schulz*** discloses the well known technique of having and retrieving a textual representation of an audio signal, obtaining a portion of the audio signal corresponding to the segment of the textual representation, and providing the segment of the textual representation and the portion of the audio signal to the user (column 5 lines 30-33, *text is synchronized with a specific spoken word during playback of an audio file*). The synchronization of audio and text is used in ***Schulz*** to improve the editing efficiency, since the synchronized text and audio speeds up the process for the user. The known technique of audio and text synchronization, as disclosed in ***Schulz***, was combined with ***Shiotani*** to improve the device in the same way, i.e. enable improved editing (translation correction) efficiency.

3. Applicant's arguments with respect to claims 20, 21 and 40 are similar to those recited above; therefore the examiner respectfully disagrees for the reasons cited above.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1-40 and 42-45 and 47 are rejected under 35 U.S.C. 103(a) as being unpatentable over ***Shiotani*** (4,814,988) in view of ***Schulz*** (6,360,237).

4. As per claims 1 and 20, ***Shiotani*** discloses a method for facilitating translation of an audio signal that includes speech to another language, comprising:

Retrieving a textual representation (column 2 lines 11-14);

Presenting the textual representation to a user (column 2 lines 14-16);

Receiving selection of a segment of the textual representation for translation (column 2 lines 16-21);

Receiving translation made by the user of the portion of the audio signal (column 2 lines 39-41, *the user provides correction of the translation result of the specified input region*).

Shiotani does not disclose having and retrieving a textual representation of an audio signal, obtaining a portion of the audio signal corresponding to the segment of the textual representation, providing the segment of the textual representation and the portion of the audio signal to the user. However, ***Schulz*** discloses that it is well known to use automatic speech recognition to convert spoken language into written text (column 1 lines 27-34), which is then further processed. In addition, ***Schulz*** discloses a system that synchronizes text with a specific spoken word during playback of an audio

file (column 5 lines 30-33). In ***Schulz***, the user processes text displayed on a screen during playback of an audio file. All of the elements of claims 1 and 20 are known in references ***Shiotani*** and ***Schulz***, the only difference is their combination for use in a machine translation system.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to retrieve a textual representation of an audio signal for translation in ***Shiotani***, since it would enable the system to translate spoken language as well as textual documents.

It would also have been obvious to one of ordinary skill in the art at the time of the invention to obtain a portion of the audio signal corresponding to the segment of the textual representation and provide the segment of the textual representation and the portion of the audio signal to the user in ***Shiotani***, since operation of the text editor is in no way dependent on the type of editing performed by the user, i.e. translation or transcription, and the combination of the text editing software with a standard machine translation system would produce the predictable result of enabling the user to quickly and easily edit, or translate, text displayed on the monitor without interruption during playback of the speech from an audio recording, as indicated in ***Schulz*** (column 5 lines 55-58).

5. As per claim 21, ***Shiotani*** discloses a translation system, comprising:

Obtaining a textual representation (column 2 lines 11-14);

Present the transcription to a user (column 2 lines 14-16);

Receive selection of a portion of the transcription for translation (column 2 lines 16-21);

Receive from the user a translation made by the user of the portion of the audio signal (column 2 lines 39-41, *the user provides correction of the translation result of the specified input region*).

Shiotani does not disclose a memory configured to store instructions, and a processor configured to execute the instructions in memory to perform the aforementioned steps as well as, obtain a transcription of an audio signal that includes speech, retrieve a portion of the audio signal corresponding to the portion of the transcription, and provide the portion of the transcription and the portion of the audio signal to the user. However, ***Shiotani*** does disclose the use of an OCR (optical character reader), a CRT screen, and an input buffer, such as in memory. These elements suggest the invention is performed on a computer system. In addition, ***Schulz*** discloses that it is well known for computer systems to include a data storage medium (memory) with a program (instructions) for performing a specific function (column 20-25). ***Schulz*** also discloses that it is well known to use automatic speech recognition to convert spoken language into written text (column 1 lines 27-34), which is then further processed. ***Schulz*** further discloses a system that synchronizes text with a specific spoken word during playback of an audio file (column 5 lines 30-33). In ***Schulz***, the user processes text displayed on

a screen during playback of an audio file. All of the elements of claim 21 are known in references ***Shiotani*** and ***Schulz***, the only difference is their combination for use in a machine translation system.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have a memory and processor configure to execute the instruction stored in memory in ***Shiotani***, since a computer system can perform calculations and execute instructions extremely quickly, thus decreasing processing time and enabling a real-time application.

It would have been obvious to one of ordinary skill in the art at the time of the invention to obtain a transcription of an audio signal that includes speech in ***Shiotani***, since it would enable the system to translate spoken language as well as textual documents.

It would also have been obvious to one of ordinary skill in the art at the time of the invention to obtain a portion of the audio signal corresponding to the segment of the textual representation and provide the segment of the textual representation and the portion of the audio signal to the user in ***Shiotani***, since operation of the text editor is in no way dependent on the type of editing performed by the user, i.e. translation or transcription, and the combination of the text editing software with a standard machine translation system would produce the predictable result of enabling the user to quickly and easily edit, or translate, text displayed on the monitor without interruption during

playback of the speech from an audio recording, as indicated in ***Schulz*** (column 5 lines 55-58).

6. As per claims 2 and 22, ***Shiotani*** in view of ***Schulz*** disclose the method of claims 1 and 21, and ***Shiotani*** does not explicitly disclose wherein the retrieving a textual representation includes generating a request for information, sending the request to a server, and obtaining, from the server, at least the textual representation of the audio signal. However, ***Shiotani*** does disclose that a textual representation of an input sentence is accessed from an input buffer, or memory, and then displayed on a screen (column 2 lines 10-18). In addition, in any computer system program instructions are executed in order to retrieve information from memory, such as a server.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to send a request for information to a server and obtain a textual representation of the audio signal in ***Shiotani***, since it would enable the system to process information previously stored in memory.

7. As per claims 3 and 23, ***Shiotani*** in view of ***Schulz*** disclose the method of claims 1 and 21, and ***Schulz*** further discloses wherein the presenting the textual representation to a user, includes: obtaining the audio signal, providing the audio signal and the textual representation of the audio signal to the user, and visually synchronizing the providing of the audio signal with the textual representation of the audio signal

(column 5 lines 30-33 and column 6 lines 29-30, *the audio signal is provided the user, synchronized with the test. Therefore the audio signal must have first been obtained*). ***Schulz*** discloses a system that synchronizes text with a specific spoken word during playback of an audio file (column 5 lines 30-33). In ***Schulz***, the user processes text displayed on a screen during playback of an audio file. All of the elements of claims 3 and 23 are known in references ***Shiotani*** and ***Schulz***, the only difference is their combination for use in a machine translation system.

Therefore it would also have been obvious to one of ordinary skill in the art at the time of the invention to obtain the audio signal, provide the audio signal and the textual representation of the audio signal to the user, and visually synchronize the providing of the audio signal with the textual representation of the audio signal in ***Shiotani***, since operation of the text editor is in no way dependent on the type of editing performed by the user, i.e. translation or transcription, and the combination of the text editing software with a standard machine translation system would produce the predictable result of enabling the user to quickly and easily edit, or translate, text displayed on the monitor without interruption during playback of the speech from an audio recording, as indicated in ***Schulz*** (column 5 lines 55-58).

8. As per claims 4 and 24, ***Shiotani*** in view of ***Schulz*** disclose the method of claims 3 and 23, and ***Schulz*** further discloses wherein the obtaining the audio signal includes accessing a database of original media to retrieve the audio signal (column 5

lines 30-33, *the audio recording is played back and aligned with the words on the screen. The audio played back is from an audio recording; therefore the audio must have been accessed from a recording medium or memory, such as a database*).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to access a database of original media to retrieve the audio signal in ***Shiotani***, since a database enables the system to store information for processing at a later time.

9. As per claims 5,8,25, and 28 ***Shiotani*** in view of ***Schulz*** disclose the method of claims 3,1,23 and 21, and ***Schulz*** further discloses wherein the obtaining the audio signal includes receiving input, from the user, regarding a desire for the audio signal (column 12 line 63-column 13 line 12, *if the user enters a command to start playback of the audio signal, the playback edit function mode is entered, otherwise the system enters the standard editing mode*) initiating a media player, and using the media player to obtain the audio signal (column 12 line 63-column 13 line 12, *if the user enters a command to start playback of the audio signal, the playback edit function mode is entered and playback of the audio recording synchronized with the text begins. Since the audio, a type of media, is output, it must be have been obtained and output through a media player*).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to receive input, from the user, regarding a desire for the audio signal,

initialize a media player, and use the media player to obtain the audio signal in ***Schulz***, since it would enable the user to choose between a standard editing or translation mode, or a playback translation or edit mode, where the audio is played back synchronized with the text.

10. As per claims 6 and 26, ***Shiotani*** in view of ***Schulz*** disclose the method of claims 1 and 21, and ***Shiotani*** does not explicitly disclose wherein the receiving selection of a segment of the textual representation includes identifying a portion of the textual representation selected by the user, accessing a server to obtain text corresponding to the portion of the textual representation, and receiving, from the server, the text corresponding to the portion of the textual representation. However, ***Shiotani*** does disclose that a textual representation of an input sentence is accessed from an input buffer, or memory, and then displayed on a screen (column 2 lines 10-18). In addition, in any computer system program instructions are executed in order to retrieve information from memory, such as a server.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to send a request for information to a server and obtain a textual representation of the audio signal in ***Shiotani***, since it would enable the system to process information previously stored in memory.

11. As per claims 7 and 27, ***Shiotani*** in view of ***Schulz*** disclose the method of claims 6 and 26, and ***Schulz*** further discloses wherein the text includes a transcription of the audio signal and metadata corresponding to the portion of the textual representation (column 4 lines 52-59, *a file containing the transcription of the input speech also contains beginning and end times for each word and silent pauses*).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have text file that includes a transcription and metadata in ***Shiotani***, since it would enable the system to locate pauses, and suppress them during playback, as indicated in ***Schulz*** (column 4 lines 60-65).

12. As per claims 9 and 29, ***Shiotani*** in view of ***Schulz*** disclose the method of claims 8 and 28, and ***Schulz*** further discloses wherein the using the media player includes identifying, by the media player, the segment of the textual representation, and retrieving the portion of the audio signal corresponding to the segment of the textual representation (column 6 lines 18-30, *the system uses the beginning and ending times of words to align the cursor on the monitor with a particular displayed word during playback of the audio recording. Since the audio is played back synchronized with the time information from the text file, a media player must have identified the textual representation and retrieved the audio signal*).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to identify, by the media player, the segment of the textual

representation, and retrieve the portion of the audio signal corresponding to the segment of the textual representation in ***Shiotani***, since operation of the text editor is in no way dependent on the type of editing performed by the user, i.e. translation or transcription, and the combination of the text editing software with a standard machine translation system would produce the predictable result of enabling the user to quickly and easily edit, or translate, text displayed on the monitor without interruption during playback of the speech from an audio recording, as indicated in ***Schulz*** (column 5 lines 55-58).

13. As per claims 10,11, 30 and 31, ***Shiotani*** in view of ***Schulz*** disclose the method of claims 9 and 29, and ***Schulz*** further discloses wherein the segment of the textual representation includes a starting position in the textual representation, and wherein the identifying the segment includes identifying a time codes associated with the beginning and ending of the textual representation (column 6 lines 18-30, *the system uses the beginning and ending times of words to align the cursor on the monitor with a particular displayed word during playback of the audio recording*).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have a textual representation that includes a starting position, and identify time codes associated with the beginning and end times of the textual representation in ***Shiotani,*** since it would enable the user to quickly and easily edit, or

translate, text displayed on the monitor without interruption during playback of the speech from an audio recording, as indicated in ***Schulz*** (column 5 lines 55-58).

14. As per claims 12 and 32, ***Shiotani*** in view of ***Schulz*** disclose the method of claims 1 and 21, and ***Shiotani*** further discloses wherein the providing the segment of the textual representation and the portion of the audio signal to the user includes displaying the segment of the textual representation in a same window as will be used by the user to provide the translation of the portion of the audio signal, including as a split screen in a translation window (column 2 lines 15-20 and Figure 4(a) and 4(b)).

15. As per claims 13 and 33, ***Shiotani*** in view of ***Schulz*** disclose the method of claims 1 and 21, and ***Schulz*** further discloses wherein the providing the segment of the textual representation and the portion of the audio signal to the user includes visually synchronizing the providing of the portion of the audio signal with the segment of the textual representation (column 5 lines 30-33 and column 6 lines 29-30). ***Schulz*** discloses a system that synchronizes text with a specific spoken word during playback of an audio file (column 5 lines 30-33). In ***Schulz***, the user edits text displayed on a screen during playback of an audio file. All of the elements of claims 13 and 33 are known in references ***Shiotani*** and ***Schulz***, the only difference is their combination for use in a machine translation system.

Therefore it would also have been obvious to one of ordinary skill in the art at the time of the invention to provide the segment of the textual representation and the portion of the audio signal to the user by visually synchronizing the providing of the portion of the audio signal with the segment of the textual representation in ***Shiotani***, since operation of the text editor is in no way dependent on the type of editing performed by the user, i.e. translation or transcription, and the combination of the text editing software with a standard machine translation system would produce the predictable result of enabling the user to quickly and easily edit, or translate, text displayed on the monitor without interruption during playback of the speech from an audio recording, as indicated in ***Schulz*** (column 5 lines 55-58).

16. As per claims 14 and 34, ***Shiotani*** in view of ***Schulz*** disclose the method of claims 13 and 33, and ***Schulz*** further discloses wherein the segment of the textual representation includes time codes corresponding to when words in the textual representation were spoken (column 4 lines 52-59, *a file containing the transcription of the input speech also contains beginning and end times for each word and silent pauses*).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have a textual representation that includes time codes corresponding to when words in the textual representation were spoken in ***Shiotani***, since it would

enable the system to locate pauses, and suppress them during playback, as indicated in ***Schulz*** (column 4 lines 60-65).

17. As per claims 15 and 35, ***Shiotani*** in view of ***Schulz*** disclose the method of claims 14 and 34, and ***Schulz*** further discloses wherein the visually synchronizing the providing of the portion of the audio signal with the segment of the textual representation includes comparing times corresponding to the providing of the portion of the audio signal to the time codes from the segment of the textual representation, and visually distinguishing words in the segment of the textual representation when the words are spoken during the providing of the portion of the audio signal (column 6 lines 18-30).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to compare times corresponding to the providing of the portion of the audio signal to the time codes from the segment of the textual representation, and visually distinguishing words in the segment of the textual representation when the words are spoken during the providing of the portion of the audio signal in ***Shiotani***, since operation of the text editor is in no way dependent on the type of editing performed by the user, i.e. translation or transcription, and the combination of the text editing software with a standard machine translation system would produce the predictable result of enabling the user to quickly and easily edit, or translate, text

displayed on the monitor without interruption during playback of the speech from an audio recording, as indicated in ***Schulz*** (column 5 lines 55-58).

18. As per claims 16,17,36 and 37, ***Shiotani*** in view of ***Schulz*** disclose the method of claims 1 and 21, and ***Schulz*** further discloses wherein the providing the segment of the textual representation and the portion of the audio signal to the user includes permitting the user to control the providing of the portion of the audio signal by allowing the user to at least one of fast forward, speed up, slow down, and back up the providing of the portion of the audio signal using foot pedals (column 2 lines 29-34).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to control the providing of the portion of the audio signal by allowing the user to at least one of fast forward, speed up, slow down, and back up the providing of the portion of the audio signal using foot pedals in ***Schulz***, in order to achieve efficient use of the various inputs and controls.

19. As per claims 18 and 38, ***Shiotani*** in view of ***Schulz*** disclose the method of claims 16 and 36, and ***Schulz*** further discloses wherein the permitting the user to control the providing of the portion of the audio signal includes permitting the user to rewind the portion of the audio signal at least one of a predetermined amount of time and a predetermined amount of words (column 2 line 29-34, *the user can use keyboard*

input or a foot control to control the audio signal, including moving forward and rewinding).

20. As per claims 19 and 39, ***Shiotani*** in view of ***Schulz*** disclose the method of claims 1 and 21, and ***Shiotani*** further discloses publishing the translation to a user-determined location (column 3 lines 2-4).

21. As per claim 40, ***Shiotani*** discloses a graphical user interface, comprising:

A text input section that includes text information in a first language (column 2 lines 11-14);

A translation section that receives a translation made by the user of the non-text information into a second language (column 2 lines 14-16);

Shiotani does not disclose transcription section that includes a transcription of non-text information in a first language, and a play button that, when selected, causes the retrieval of the non-text information to be initiated, playing of the non-text information, and the playing of the non-text information to be visually synchronized with the transcription in the transcription section. ***Schulz*** discloses that it is well known to use automatic speech recognition to convert spoken language into written text, i.e. a transcript (column 1 lines 27-34), which is then further processed. In addition, ***Schulz*** discloses a system that synchronizes text with a specific spoken word during playback

of an audio file (column 5 lines 30-33), and indicates that controls for the text viewer and audio play back include a keyboard or foot pedal (play button) (column 2 lines 29-32). In ***Schulz***, the user processes text displayed on a screen during playback of an audio file. All of the elements of claim 40 are known in references ***Shiotani*** and ***Schulz***, the only difference is their combination for use in a machine translation system.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have transcription section that includes a transcription of non-text information in a first language in ***Shiotani***, since it would enable the system to translate spoken language as well as textual documents.

It would also have been obvious to one of ordinary skill in the art at the time of the invention to a have a play button that, when selected, causes the retrieval of the non-text information to be initiated, playing of the non-text information, and the playing of the non-text information to be visually synchronized with the transcription in the transcription section in ***Shiotani***, since operation of the text editor is in no way dependent on the type of editing performed by the user, i.e. translation or transcription, and the combination of the text editing software with a standard machine translation system would produce the predictable result of enabling the user to quickly and easily edit, or translate, text displayed on the monitor without interruption during playback of the speech from an audio recording, as indicated in ***Schulz*** (column 5 lines 55-58).

22. As per claims 42 and 43, ***Shiotani*** in view of ***Schulz*** disclose the graphical user interface of claim 40, but neither ***Shiotani*** nor ***Schulz*** explicitly disclose a configuration button, that when selected, causes a window to be presented, the window permitting an amount of backup to be specified, the amount of backup including one of a predetermined amount of time and a predetermined number of words, and wherein the window further permits a name to be given for the translation and a location of publication to be specified. However, ***Shiotani*** does disclose a translation buffer for storing the result of translation of a selected portion of the input (column 2 lines 38-41). The translation buffer stores a predetermined number of words, i.e. the region of the text specified by the user and then translated. In addition, the use of a configuration button to present a window that permits a name to be given to a file and a location of publication to be specified is a feature of any text editing or word processing software, running on any of a number of operating systems, such as windows and Linux. The software enables the user to use the save button (configuration button), located under a file menu in a task bar, to choose a location in memory as well as a name for the file.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have a configuration button, that when selected, causes a window to be presented, the window permitting an amount of backup to be specified, the amount of backup including one of a predetermined amount of time and a predetermined number of words, and wherein the window further permits a name to be given for the translation and a location of publication to be specified in ***Shiotani***, since it would

enable the system to save the file in memory so that it can be easily retrieved for further processing in the future.

23. As per claim 44, ***Shiotani*** in view of ***Schulz*** disclose the graphical user interface of claim 40, and ***Schulz*** further discloses wherein the play button further causes words in the transcription to be visually distinguished in synchronism with the words in the non-text information being played (column 6 lines 18-30).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have a play button that causes words in the transcription to be visually distinguished in synchronism with the words in the non-text information being played in ***Shiotani***, since operation of the text editor is in no way dependent on the type of editing performed by the user, i.e. translation or transcription, and the combination of the text editing software with a standard machine translation system would produce the predictable result of enabling the user to quickly and easily edit, or translate, text displayed on the monitor without interruption during playback of the speech from an audio recording, as indicated in ***Schulz*** (column 5 lines 55-58).

24. As per claim 45, ***Shiotani*** in view of ***Schulz*** disclose the graphical user interface of claim 40, and ***Schulz*** further discloses wherein the non-text information includes at

least one of audio and video (column 4 lines 46-59, *a speech recognition unit converts a recording of speech (audio non-text information) into a text file*).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to process non-text information that includes at least one of audio and video in ***Shiotani***, since it would enable the system to translate spoken language as well as textual documents.

25. As per claim 47, ***Shiotani*** discloses a method comprising:

A user viewing a textual transcription of information in a first language on a transcription section (column 2 lines 14-16 and Figure 4(a) and 4(b));

Said user translating said information thereby obtaining a translation in a second language, said user using a different section of said graphical user interface (GUI) to display said translation while making said translation (column 2 lines 16-21, and Figures 4(a) and 4(b), whereby the synchronizing of said audio playback with said textual transcription aids said user in making said translation.

Shiotani does not disclose a user listening to an audio playback of information in a first language while viewing a textual transcription of said information in said first language on a transcription section of a graphical user interface (GUI), said textual transcription being synchronized with said audio playback, said user translating the audio playback of said information. In addition, ***Schulz*** discloses a system that synchronizes text with a

specific spoken word during playback of an audio file (column 5 lines 30-33). In ***Schulz***, the user processes text displayed on a screen during playback of an audio file. All of the elements of claims 1 and 20 are known in references ***Shiotani*** and ***Schulz***, the only difference is their combination for use in a machine translation system.

Therefore it would also have been obvious to one of ordinary skill in the art at the time of the invention to enable a user to listen to an audio playback of information in a first language while viewing a textual transcription of said information in said first language on a transcription section of a graphical user interface (GUI), said textual transcription being synchronized with said audio playback, said user translating the audio playback of said information in ***Shiotani***, since the combination of the known text and audio synchronization technique with a standard machine translation system would produce the predictable result of enabling the user to quickly and easily edit, or translate, text displayed on the monitor without interruption during playback of the speech from an audio recording, as indicated in ***Schulz*** (column 5 lines 55-58).

Claims 41 and 46 are rejected under 35 U.S.C. 103(a) as being unpatentable over ***Shiotani*** in view of ***Schulz*** as applied to claim 40 above, and further in view of ***Saindon*** (6,820,055).

26. ***Shiotani*** in view of ***Schulz*** disclose the graphical user interface of claim 40, however neither disclose wherein the transcription visually distinguishes names of

people, places, and organizations and wherein the graphical user interface is associated with a word processing application. ***Saindon*** discloses a system for automated transcription and translation that processes text to visually distinguish the names of people, places and organizations using a word processor (column 16 lines 34-65, *the system processes the text to determine if all proper nouns are capitalized using software such as Microsoft word*).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have a transcription that visually distinguishes names of people, places, and organizations and a graphical user interface is associated with a word processing application in ***Shiotani*** and ***Schulz***, since it would enable the system to generate text that provides accurate translations, as indicated in ***Saindon*** (column 16 lines 38-40), using reliable commercially established software that is readily available.

Conclusion

THIS ACTION IS MADE FINAL. Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dorothy Sarah Siedler whose telephone number is 571-270-1067. The examiner can normally be reached on Mon-Thur 9:30am-5:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 571-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

DSS

RICHEMOND DORVIL
SUPERVISORY PATENT EXAMINER